

学校编码: 10384

分类号 _____ 密级 _____

学号: 23020091152699

UDC _____

厦 门 大 学

专 业 硕 士 学 位 论 文

基于 Oracle 海量数据库的 SQL 优化研究

Research on the Technology of SQL Tuning
Based on Oracle Huge Database

陈贻鑫

指导教师姓名: 倪子伟 副教授

专 业 名 称: 计算机技术

论文提交日期: 2012 年 5 月

论文答辩时间: 2012 年 6 月

学位授予日期:

答辩委员会主席: _____

评 阅 人: _____

2012 年 5 月

厦门大学博硕士论文摘要库

厦门大学学位论文原创性声明

本人呈交的学位论文是本人在导师指导下,独立完成的研究成果。本人在论文写作中参考其他个人或集体已经发表的研究成果,均在文中以适当方式明确标明,并符合法律规范和《厦门大学研究生学术活动规范(试行)》。

另外,该学位论文为()课题(组)的研究成果,获得()课题(组)经费或实验室的资助,在()实验室完成。(请在以上括号内填写课题或课题组负责人或实验室名称,未有此项声明内容的,可以不作特别声明。)

声明人(签名):

年 月 日

厦门大学博硕士论文摘要库

厦门大学学位论文著作权使用声明

本人同意厦门大学根据《中华人民共和国学位条例暂行实施办法》等规定保留和使用此学位论文，并向主管部门或其指定机构送交学位论文(包括纸质版和电子版)，允许学位论文进入厦门大学图书馆及其数据库被查阅、借阅。本人同意厦门大学将学位论文加入全国博士、硕士学位论文共建单位数据库进行检索，将学位论文的标题和摘要汇编出版，采用影印、缩印或者其它方式合理复制学位论文。

本学位论文属于：

()1. 经厦门大学保密委员会审查核定的保密学位论文，于
年 月 日解密，解密后适用上述授权。

()2. 不保密，适用上述授权。

(请在以上相应括号内打“√”或填上相应内容。保密学位论文应是已经厦门大学保密委员会审定过的学位论文，未经厦门大学保密委员会审定的学位论文均为公开学位论文。此声明栏不填写的，默认为公开学位论文，均适用上述授权。)

声明人(签名)：

年 月 日

厦门大学博硕士论文摘要库

摘 要

随着网络的普及和电子商务的不断发展,以及数据库规模的不断扩大、信息量不断增加,对数据库的性能要求也不断提高,系统的响应速度就成为目前系统需要解决的最主要的问题之一。同时,随着信息化进程的加快,以前仍属于小型的数据库,现在也变成了海量数据库;以前没有受到太多关注的数据,现在也备受关注。因此,作为数据最终载体的数据库,其性能直接影响了整个应用系统的性能,而数据库的响应速度,则是数据库性能的一个重要指标。

性能优化是数据库应用的核心问题。目前的数据库产品发展已日臻成熟,数据库管理员除了日常的常规维护工作之外,大多精力都花在对数据库的优化上面。而数据库的优化中,SQL 优化是极其重要的一环,这是因为对于海量数据库,两个等价的 SQL 语句之间的执行速度却可以存在着数倍甚至上百倍的差别,如何从众多等价 SQL 语句中选择执行效率最高的优质 SQL 语句和执行计划,直接对数据库查询速度有着重大的影响,由此可见 SQL 语句的优化是数据库优化的一个重要的部分,也是数据库管理员日常工作的一大课题。

本文以目前使用最为广泛、技术最为成熟的 Oracle 10g 版本数据库为平台,以 SQL 语句为研究主体,结合 Oracle 10g 新特性,对如何有效优化 SQL 进行深入研究。

论文的内容主要包括以下几个方面:

1. 对 Oracle 10g 的体系结构进行整体介绍,阐述各个部分在 SQL 执行中所起的作用。
2. 对 SQL 在 Oracle 中的执行过程和解析顺序进行详细分析。
3. 对比等价 SQL 语句在执行时的差异,总结优质 SQL 语句的书写思路。
4. 研究常用 SQL 优化方法,结合 Oracle 10g 的新优化工具,总结出一套 SQL 优化的系统方案,并通过具体实验案例,表明其在 SQL 优化方面的有效性。

关键词: Oracle; SQL 优化; 数据查询

厦门大学博硕士论文摘要库

Abstract

With the development of popularization of network and e-business, the size of database is constantly expanding, and the content of information continues to increase, database performance requirement is also rising. Response speed of the database system becomes one of the main problems of the current system needed to be addressed. At the same time, as the speeding up of the process of information, the database which belonged to the small before turned into a massive database now; the data which has not been paid much attention before is received much concern now. As the ultimate vehicle of data -- database, its performance directly affects the performance of the entire system and database response time is an important indicator of the performance of the database.

Performance tuning is the key problem of database application problems. Current database product development is increasingly more sophisticated. Database administrators spent most of their energies on database optimization, in addition to daily regular maintenance work. SQL tuning is a very important part in database optimization. This is because, for huge database, the execution speeds between two equivalent statements in SQL can be a few times even hundred times of differences. How to choose the most efficient SQL statement and explain plan in the equivalent SQL statements as high quality SQL statements had a significant impact on database query speed directly, so SQL tuning is an important part of database optimization and is also a major issue of the routine work of database administrators.

Based on Oracle 10g database which version currently is the most widely used and has most sophisticated technology and taking SQL statements as research subjects, this paper combines the new features of Oracle 10g and makes in-depth research on how to effectively SQL tuning.

The paper mainly includes the following aspects:

1. Overview to the Oracle 10g architecture, explained the role of the various parts in SQL execution.
2. Detailed discussion the process of SQL execution and the order of analyze in Oracle.

3. Comparison of the difference of equivalent SQL statement execution, sum up general thought of writing the high quality of SQL statements.

4. Research on common method of SQL tuning, combined with new optimization tools of Oracle 10g, came to an Effective SQL tuning scheme, and shown its effectiveness in SQL tuning by some specific test cases.

Key words: Oracle 10g; SQL Tuning; Data Query

目录

第一章 绪论.....	1
1.1 课题研究背景及意义.....	1
1.2 国内外研究现状.....	2
1.3 本文的主要研究内容及章节安排.....	3
第二章 Oracle 10g 体系结构及调优组成.....	4
2.1 Oracle 10g 体系结构	4
2.1.1 Oracle 10g 体系结构概述.....	4
2.1.2 Oracle 内存结构.....	5
2.1.3 Oracle 物理结构.....	9
2.2 Oracle 10g 工作流程	10
2.3 Oracle 10g 优化层次	12
2.4 Oracle 10g 优化的性能收益	13
2.5 本章小结.....	14
第三章 SQL 查询语句分析	15
3.1 SQL 语句的解析过程.....	15
3.2 SQL 的连接访问方式.....	17
3.2.1 SQL 的连接方式.....	17
3.2.2 SQL 的访问方式.....	19
3.3 等价 SQL 的性能差异.....	22
3.4 本章小结.....	30
第四章 基于 Oracle10g 新特性的 SQL 优化研究	31
4.1 定位问题 SQL.....	31
4.1.1 Statspack.....	32
4.1.2 自动工作量资料档案库	32
4.1.3 数据库自动诊断监视程序	35
4.1.4 TOP SQL	37
4.1.5 动态性能视图	38
4.2 分析问题 SQL.....	38

4.2.1 查询优化器	38
4.2.2 执行计划	44
4.3 提出优化措施.....	48
4.3.1 自动调整优化器	48
4.3.2 SQL 优化指导.....	49
4.3.3 SQL 访问指导.....	52
4.4 本章小结.....	54
第五章 Oracle SQL 优化实验与分析.....	55
5.1 实例背景.....	55
5.2 优化案例.....	56
5.2.1 优化案例一：大数据量表连接顺序优化	56
5.2.2 优化案例二：返回结果集小的大数据量表连接方式优化	61
5.2.3 优化案例三：组合查询条件过滤性强的大数据量表访问方式优化	64
5.2.4 优化案例四：带有大量计算的大数据量表连接方式优化	67
5.3 本章小结.....	71
第六章 总结与展望	72
参考文献.....	74
致谢.....	76

Contents

Chapter 1.Introduction.....	1
1.1 Background and Significance.....	1
1.2 Research Stauts at Home and Abroad	2
1.3 Main Content and Arrangement of the Thesis.....	3
Chapter 2.Oracle10g Architecture and the Component of Tuning.....	4
2.1 Oracle 10g Architecture.....	4
2.1.1 Overview	4
2.1.2 Oracle Memory Structure	5
2.1.3 Oracle Physical Structure	9
2.2 Oracle 10g Work Process	10
2.3 Oracle 10g Tuning Layer.....	12
2.4 Performance Improvement of Oracle 10g Tuning	13
2.5 Summary	14
Chapter 3.Analysis of SQL Query Statement.....	15
3.1 SQL Statement Processing Phase	15
3.2 SQL Access and Join Path.....	17
3.2.1 SQL Access Path	17
3.2.2 SQL Join Path.....	19
3.3 Performance Difference of Equivalence SQL Statements.....	22
3.4 Summary	30
Chapter 4.SQL Tuning Based on Oracle 10g New Features.....	31
4.1 Identification of High-Load SQL.....	31
4.1.1 Statspack.....	32
4.1.2 AWR.....	32
4.1.3 ADDM.....	35
4.1.4 TOP SQL	37
4.1.5 Dynamic Performance View	38
4.2 Analysis of High-Load SQL	38

4.2.1	Query Optimizer.....	38
4.2.2	Explanation Plan.....	44
4.3	Take Corrective Action.....	48
4.3.1	Auto Tuning Optimizer	48
4.3.2	SQL Tuning Advisor	49
4.3.3	SQL Access Advisor.....	52
4.4	Summary	54
Chapter 5.Experiments and Analysis of Oracle SQL Tuning		55
5.1	Experiment Background	55
5.2	Tuning Case	56
5.2.1	Case1: Huge Data Table Join Order Tuning.....	56
5.2.2	Case2: Huge Data Table Join Method Tuning	61
5.2.3	Case3: Table Access Method Tuning with Strong Filterability	64
5.2.4	Case4: Huge Data Table Access Method Tuning with Massive Calculation	67
5.3	Summary	71
Charper 6.Conclusion and Prospect		72
References		74
Acknowledgements		76

第一章 绪论

1.1 课题研究背景及意义

随着计算机的普及和信息网络用户爆炸性地增长,网络信息量呈几何增长,人们对数据信息的需求直接反映在对数据库系统上面。为了应对这样一个信息爆炸的时代,许多企业纷纷采用海量数据库系统进行数据的存储。另一方面人们对数据库性能的要求已不仅仅满足于能够得到正确的结果而已,如何处理好这些海量数据的查询操作,以减少系统的响应时间,使得用户在更短的时间能获得想要的结果,成为企业的一大课题。

Oracle 数据库作为企业级应用的大型数据库,以其稳定和高效等诸多优点备受企业青睐,成为国内高端数据库市场的主流产品和众多行业信息化系统的主要支柱,其应用领域越来越广泛。企业资源计划系统 ERP、办公自动化系统 OA、管理信息系统 MIS、客户关系管理系统 CRM、电子商务等海量数据库应用中,都将 Oracle 数据库列为首选。因此对提高海量数据库查询响应时间的研究,自然地以 Oracle 数据库作为首选研究对象进行。而在版本的选择上,选择目前大型企业应用最为广泛、技术最为成熟的 Oracle 10g 版本作为研究平台,这符合现代应用需求,具有重要的现实意义。

结构化查询语言 SQL,是一种数据库查询和程序设计语言,它用于存取数据以及查询、更新和管理关系数据库系统,用户只需指定想要读取的字段、表以及数据的条件即可完成查询。SQL 作为数据库的访问语言,自 1970 年代诞生以来,就被广泛接受并一直流行至今。然而 SQL 语言的简单性却但带有欺骗性,由于 SQL 与生俱来的强大能力,同一查询在书写形式可以有多种形式。SQL 语句越是复杂,其组合越多,甚至趋近无限,导致人们在编写 SQL 时往往会陷入忽略不同实现方法之间存在性能差异的盲区,这种性能差异在海量数据库中尤为明显。实际上,实现同一功能的等价语句在执行效率上并不等效,甚至存在着极大差异,劣质 SQL 语句和优质 SQL 语句之间的速度差别可以达到上百倍。因此,一个劣质 SQL 语句、不适当的索引和连接方式就可能要耗费大量的系统资源,这也就意味着 SQL 优化的必要性和重要性。随着信息化的发展,数据量爆发式增长,人们不再只满足得到正确结果,而对响应速度有了更进一步要求。如何从众多等价 SQL 语句中分辨优质和劣质的 SQL 语句,优化 SQL 的执行计划,尽可能缩小数据库查询响应时间,成为了 DBA 们进行数据库优化研究的一大重点。

综上所述, 如何充分利用 Oracle 数据库的功能特性提高数据查询的响应时间已成为 Oracle 数据库管理员日常的一大课题, 而 Oracle SQL 优化在所有 Oracle 优化中占了很重要的一部分, 对于那些每天都要执行频繁查询的海量 Oracle 数据库而言, 查询语句效率的高低直接影响着响应速度, 进行恰当的 SQL 语句优化能够在很大程度上提高整个 Oracle 数据库的性能, 同时也能延长数据库服务器硬件的寿命, 而对 SQL 语句进行优化的研究, 可以为如何改进数据库设计方案、降低数据库的运行和维护成本、促进数据库技术发展提供理论依据。因此, 如何才能有效地优化 SQL, 达到数据库的理想性能, 这是一个非常值得研究的课题。

1.2 国内外研究现状

查询响应速度随着信息化社会的发展, 变得越来越重要的性能指标, 数据库的优化也越来越受到重视, 大型企业的海量数据库均有专人进行数据库维护和优化, SQL 优化的研究也随着数据库优化这个大课题而备受关注, 出现不少关于研究 SQL 优化的著作, 譬如 Oracle 法国的第一个性能优化课程编写者 Stephane Faroult 以《孙子兵法》的智慧来探讨 SQL 优化问题的《SQL 语言艺术》以及由 Oracle 公司授权、美国 Donald K.Burleson 所编写的《Oracle 高性能 SQL 调整》等等。

Oracle 非常注重 SQL 优化的研究, 随着版本的更新不断推出新的优化工具。Oracle 7 之前版本主要的优化工具是基于规则的优化器 RBO, 后来逐渐转向基于成本的优化器 CBO, 随后的 Oracle 9i 版本又提供了一系列的动态性能视图和 Statspack 报表。到了 Oracle 10g 版本则推出基于 WEB 的 Oracle Enterprise Manager 平台, 配有自动工作负载信息库 AWR 等性能信息收集统计工具和 SQL Tuning Advisor、SQL Access Advisor 两个新的 SQL 调优工具, 能通过对统计数据、访问路径、SQL 语法结构等进行分析, 给 DBA 提供 SQL 调优的建议。Oracle 11g 版本则提供自学功能进行 SQL 自动优化。

国内对于 Oracle 数据库的研究起步较晚, 2004 年时国内关于 Oracle 的原创作品还十分有限, 后来越来越多的技术爱好者开始总结、写作和分享, 关于 SQL 优化技术研究资料才开始丰富起来。期间, 这些分享和讨论主要集中在国内最大的数据库技术讨论社区 ITPUB。许多数据库专家和第一线工作者均参与其中分享自己多年优化经验讨论各种优化问题, 从索引、物化视图、优化器、书写方法等不同方面探讨研究 SQL 优化问题。即便如此, 直至 2010 年, 在国内关于数据库方面的书籍中, 写性能优化的书

Degree papers are in the "[Xiamen University Electronic Theses and Dissertations Database](#)". Full texts are available in the following ways:

1. If your library is a CALIS member libraries, please log on <http://etd.calis.edu.cn/> and submit requests online, or consult the interlibrary loan department in your library.
2. For users of non-CALIS member libraries, please mail to etd@xmu.edu.cn for delivery details.

厦门大学博硕士论文摘要库